Grade 12 Language Arts Achievement Testing Program

Provincial Report

December, 1982

Distribution:

Officials of Alberta Education
School Principals
Superintendents of Schools
The Alberta Teachers' Association
Alberta School Trustees' Association
General Public upon Request

Alberta
EDUCATION

TABLE OF CONTENTS

	Page
LIST OF TABLES	iii
LIST OF FIGURES	v
ACKNOWLEDGMENTS	vii
EXECUTIVE SUMMARY	ix
Chapter I: THE ACHIEVEMENT TESTING PROGRAM	1
Exemptions from the Achievement Testing Program	2
Chapter II: TEST DEVELOPMENT, DESCRIPTION, AND DESIGN	3
Test Development	3
Test Description	4
Part A - Written Expression	4
Part B - Multiple-Choice	5
Determination of the Student Population	14
Administration	14
Data Collection	15
Test Design Statistics	15
Chapter III: SCORING OF WRITTEN EXPRESSION	17
Chapter IV: RESULTS	21
Guide to Interpretation	21
Absentees	24
Part A - Written Expression Results	25
Part B - Multiple-Choice Results	28
Chapter V: CONCLUDING OBSERVATIONS	33

LIST OF TABLES

		Page
Table 1:	Distribution of Passage Types	5
Table 2:	English 30 Item Distribution	12
Table 3:	English 33 Item Distribution	13
Table 4:	Frequencies of Total Expression Scores . . .	26
Table 5:	Frequencies of Thought and Detail Scores . .	26
Table 6:	Frequencies of Organization Scores	26
Table 7:	Frequencies of Matters of Choice Scores . . .	27
Table 8:	Frequencies of Matters of Convention Scores .	27
Table 9:	English 30 Multiple-Choice Results	28
Table 10:	Provincial Norms - English 30	29
Table 11:	English 33 Multiple-Choice Results	30
Table 12:	Provincial Norms - English 33	31

LIST OF FIGURES

		Page
Figure 1:	Sample Passage - English 30	7
Figure 2:	Sample Items - English 30	8
Figure 3:	Sample Passage - English 33	9
Figure 4:	Sample Items - English 33	10
Figure 5:	The Written Expression Scoring Guide . . .	19

ACKNOWLEDGMENTS

The successful administration of the English 30 and English 33 Achievement Tests was the result of the concerted efforts of all involved. Success would have been impossible without substantial contributions from many people, particularly the students, teachers, and administrators who extended their full co-operation.

The technical expertise of Dr. T. O. Maguire, Professor, Division of Educational Research Services, University of Alberta, has been particularly valuable in this first application of the Achievement Testing Program. His contributions to the design, analysis, and reporting are gratefully acknowledged.

Lloyd E. Symyrozum
Director
Student Evaluation Branch

EXECUTIVE SUMMARY

This report presents the provincial results of the English 30 and English 33 Achievement Tests administered June 8, 1982, as part of the Student Achievement Testing Program of Alberta Education. Nearly 7500 students wrote the English 30 test and nearly 3800 students wrote the English 33 test. This report is intended to provide more information about the tests, the test development process, and the results for both the written expression and the multiple-choice sections. It is written, also, to assist local boards and administrators in interpreting the results for their jurisdictions.

English 30 Findings:

In Part A - Written Expression, 83% of the students produced compositions that were judged to be competent or better.

In Part B - Multiple-Choice, the provincial average score was 54%.

English 33 Findings:

In Part A - Written Expression, 37% of the students produced compositions that were judged to be competent or better.

In Part B - Multiple-Choice, the provincial average score was 44%.

Combined Grade 12 Results for the Written Expression Section:

Part A - Written Expression was common to both English 30 and English 33 tests. The combined results were as follows: 67% of the students produced compositions that were judged to be competent or better, 30% of the compositions showed limited achievement, and 3% were unacceptable.

Chapter I

THE ACHIEVEMENT TESTING PROGRAM

The Achievement Testing Program is designed to provide extensive and reliable indicators of levels of student achievement in basic education throughout the province. The tests are specific to the provincial *Program of Studies* prescribed by the Minister of Education. The curriculum specifications, which identify the instructional emphases to be placed on the content and objectives of the curriculum, serve as the basis for each of the achievement tests. These tests provide data that facilitate the establishment of provincial standards for the long-term appraisal of student achievement.

The program tests levels of student achievement in science, mathematics, social studies, and the language arts in Grades 3, 6, 9, and 12.* Not all tests are given in each grade each year; rather, the subject areas are tested on a cyclical basis. In 1982, achievement tests were administered in Grade 3 mathematics, Grade 6 science, and Grade 12 language arts. The Grade 9 social studies test was postponed because the new program had not been fully implemented across the province. The provincially set achievement tests are designed to yield results meaningful at both the jurisdictional and provincial levels.

Following each test administration in June, the results are reported in the fall of the following school year. A Summary Report that presents the highlights of the test administration is prepared for public distribution. This report, the Provincial Report, is more extensive. It is designed to assist jurisdictions in interpreting the results. Each jurisdiction also receives summary tables for that jurisdiction, including results listed by school and student. Individual statements of results are not issued to students by Alberta Education.

---

* Commencing in the 1983-84 school year, achievement testing at the Grade 12 level will be replaced by testing in Grade 11.

Exemptions from the Achievement Testing Program

Under normal circumstances, the following classes are exempt from achievement testing:

- Special Education Classes registered for grants with the Special Educational Services Branch
- classes in which the language of instruction is other than English, under the terms of Section 159 of the *School Act*
- classes in which the subject being tested has been cycled and taught in an alternate year
- classes in which the subject has been taught in a semester other than that in which the test is being administered
- classes for students for whom English is a second language

Chapter II

TEST DEVELOPMENT, DESCRIPTION, AND DESIGN

Test Development

There were three stages in the development of the English 30 and English 33 Achievement Tests: preparation of curriculum specifications, development of questions, and selection of questions for the final test.

The Curriculum Branch of Alberta Education prepared curriculum specifications that identified the major content areas, the specific objectives within each area, and the emphasis each is to receive in the classroom. The curriculum specifications were distributed to all school jurisdictions in the province, and teachers were invited to react to them.

Under the supervision of the Student Evaluation Branch of Alberta Education, test questions were written, field tested, and revised by classroom teachers. Questions were designed to reflect the reading and writing objectives outlined in the curriculum specifications. Although speaking, listening, and viewing are included in the specifications, these curricular areas were not tested.

The final test was constructed from those questions that were approved by the teacher committees. The final draft was examined by a Technical Review Committee consisting of one teacher from the subject area being tested, representatives from several stakeholder groups (the Alberta Teachers' Association, Conference of Alberta School Superintendents, and universities), and representatives of the Field Services, Curriculum, and Student Evaluation Branches of Alberta Education. The Technical Review Committee checked the draft for content validity, accuracy, and technical merit, and changes recommended by the committee were made.

Test Description

The Language Arts Achievement Tests for English 30 and English 33 each had two parts: Part A - Written Expression and Part B - Multiple-Choice. Part A - Written Expression was common to both English 30 and English 33.

Part A - Written Expression

Students were allotted two and one-half hours to write a prose composition based on one of the five statements below.

1. In a threatened world, we do possess those qualities which will enable us to survive and progress toward a better world.
2. Some part of the drama of life, e.g., conflict, chance, or coincidence, provides the raw material from which report, story, or song is created.
3. Cooperation and concern for others are important to members of any society. The non-conformist, however, may also contribute to the improvement of society.
4. In our society, a person is often identified by the work he performs. In a society without work, a new way of determining a person's identity would be needed.
5. The sting of disappointment and defeat can be lessened through humor, patience, and optimism.

Students were allowed to use a dictionary and a thesaurus, and were given suggestions for writing. Space was provided for planning activities and for rough work.

Part B - Multiple-Choice Tests

Two separate multiple-choice tests were developed for the Grade 12 Language Arts Achievement Tests, one for English 30 students and one for English 33 students. The multiple-choice tests, each consisting of 80 items, essentially examined students' understanding of "what is written" and "how it is written" in selected passages from drama, fiction, non-fiction, and poetry. Students were required to read the passages and select the best answer in each multiple-choice item. Two hours were allowed for this section of the test.

Table 1 presents information on the numbers and types of passages on each test.

Table 1

Distribution of Passage Types

	English 30 Test		English 33 Test	
	Number of Passages	Number of Items	Number of Passages	Number of Items
Drama				
Shakespearean	2	13	0	0
Modern	2	13	2	13
Fiction	3	26	1	14
Non-Fiction	2	14	4	40
Poetry	2	14	2	13
Totals	11	80	9	80

The chart indicates that the English 33 test was different from the English 30 test in that no passages from Shakespearean drama appeared in the English 33 test and a greater emphasis was placed on non-fiction passages in the English 33 test. Also, in terms of the total length of the passages in the two tests, the English 33 test had 13% fewer lines of reading than the English 30 test.

The items on the two multiple-choice tests can be separated into three major classifications or subsections:

- Meanings
- Techniques, Structures, and Forms
- Presentation of Character

Sample passages and items for each subsection are given in Figures 1 to 4.

Figure 1

Sample Passage - English 30

Mrs. Dexter: There's a big piece about it in this week's *Time*. Did you see it? *(Drawing it out of her bag.)*

Mrs. Stanley: No -- really?

Mrs. McCutcheon: Your name's in it too, Daisy. It tells all about the whole thing. Listen: "Portly Sheridan Whiteside, critic, lecturer, wit, radio orator, intimate friend of the great and near great, last week found his celebrated wit no weapon with which to combat a fractured hip. The Falstaffian Mr. Whiteside, trekking across the country on one of his annual lecture tours, met his Waterloo in the shape of a small piece of ice on the doorstep of Mr. and Mrs. Ernest W. Stanley, of Mesalia, Ohio. Result: Cancelled lectures and disappointment to thousands of adoring clubwomen in Omaha, Denver, and points west. Further result: The idol of the air waves rests until further notice in home of surprised Mr. and Mrs. Stanley. Possibility: Christmas may be postponed this year." What's *that* mean?

Mrs. Stanley: Why, what do you think of that?
(She takes the magazine; reads.)
"A small piece of ice on the doorstep of Mr. and Mrs. Ernest" -- think of it!

Mrs. McCutcheon: Of course if it were *my* house, Daisy, I'd have a bronze plate put on the step, right where he fell.

Mrs. Stanley: Well, of course I felt terrible about it. He just never goes to dinner anywhere, and he finally agreed to come here, and then *this* had to happen. Poor Mr. Whiteside! But it's going to be so wonderful having him with us, even for a little while. Just think of it! We'll sit around in the evening and discuss books and plays, all the great people he's known. And he'll talk in that wonderful way of his. . . .

THE MAN WHO CAME TO DINNER, George S. Kaufman and Moss Hart

Figure 2

Sample Items - English 30

SUBSECTION	ITEM	PERCENTAGE OF STUDENTS RESPONDING
Meanings	"... Christmas may be postponed this year" (line 14) suggests that	
	A. other towns might miss Whiteside's lecture too	22
	B. Whiteside has lost his celebrated wit	9
	*C. celebrations are not in order while Whiteside is ailing	59
	D. Whiteside doubles as Santa Claus	10
Techniques, Structures, and Forms	The exposition in this passage is provided by having	
	*A. *Mrs. McCutcheon* read aloud from the magazine	55
	B. *Mrs. Stanley* repeat the information about the piece of ice	12
	C. *Mrs. Stanley* summarize the information	16
	D. *Mrs. Dexter* introduce the topic	17
Presentation of Character	The tone of the magazine article is	
	A. objective	19
	B. approving	10
	*C. tongue-in-cheek	47
	D. derogatory	24

* *Asterisk denotes correct answer.*

Figure 3

Sample Passage - English 33

In the spring of 1935, when I was an undersized youth of fifteen, I made my first journey into arctic lands, under the tutelage of a great-uncle who was an amateur but fanatical student of birds. My Uncle Frank's consuming interest in wild things had been perpetuated in me, for from the age of six I too had been passionately interested in all the animals that haunted the rolling prairies near my home in Saskatoon. Our house in that western city had for long years been shared with pet skunks, coyotes, crows, gophers and rattlesnakes of uncertain disposition. Even my human friends were chosen almost exclusively from the half-wild and suitably ragged children at the Dundurn Indian Reserve, for I shared with those sons of nomads a strange devotion to the illusory freedom of the broad prairie plains. That restless longing to find affinity with primordial things was a legacy from my father, but it took shape and gained direction under the influence of Great-uncle Frank. He, at my mother's request, undertook to take me with him on one of his yearly pilgrimages to the ancient tundras of the arctic, where we were to spend a summer among the curious Northern birds whose very names were mystery in my ears, and whose ways of life were old before ours were begun.

In the first week of May 1935, the meadowlarks brought spring to Saskatoon -- and Uncle Frank was close behind. Tall, gaunt and weather-worn, he was still filled with the indestructible energy which had carried him through arduous years of struggle with the tough sod of an Alberta wheat farm. He had acquitted himself well in that long battle against hail and blight, and now in his late sixties he was able to indulge his lifelong hobby, and so each new spring saw him voyaging to some distant place to watch birds whose lives he longed to know and understand.

To my young eyes, Uncle Frank was a somewhat dusty Olympian and beside him I felt as insignificant as a blade of twitch grass. Yet when he looked down upon me from his great, spare height, he seemed to be vaguely satisfied with what he saw and so, on the day following his arrival at Saskatoon, we two set off together for Winnipeg, the central gateway to the arctic lands. It was the beginning of an Odyssey. . . .

PEOPLE OF THE DEER, Farley Mowat

Figure 4

Sample Items - English 33

SUBSECTION	ITEM	PERCENTAGE OF STUDENTS RESPONDING
Meanings	Mowat's interest in nature was molded by his	
	A. father	9
	B. Indian friends	15
	*C. Great-uncle Frank	57
	D. unusual pets	19
Techniques, Structures, and Forms	In line 34, the author uses allusion to	
	*A. foreshadow events to follow	35
	B. contrast the past with the present	25
	C. indicate the universal significance of the journey	30
	D. suggest a yearly pilgrimage	10
Presentation of Character	The author's attitude toward the Arctic Indians and Inuit is probably one of	
	A. condescension	16
	*B. affinity	32
	C. uncertainty	30
	D. indulgence	22

* *Asterisk denotes correct response.*

The English 30 multiple-choice section included 37 items that had previously appeared on Grade 12 examinations, 42 items that had been field tested, and one new item. The English 33 multiple-choice section was made up of 55 items that had previously appeared on Grade 12 examinations and 25 items that had been field tested. A field test is an instrument administered for the purpose of obtaining statistical information about the characteristics of individual test items.

In the English 30 test, more items appear in each of the two categories, application and synthesis, and appreciation and evaluation, than in the knowledge and understanding category. In the English 33 test, the largest number of items appear in the application and synthesis category. The English 33 test has more items in the knowledge and understanding category than does the English 30 test. The two subsections dealing with meanings have the largest number of items in both English 30 and English 33 tests.

Tables 2 and 3 show the taxonomic and subsection distributions of items for the two tests.

Table 2

English 30 Item Distribution

PART B: MULTIPLE-CHOICE

English 30 Achievement Test June 1982	Knowledge and Understanding	Application and Synthesis	Appreciation and Evaluation	Total
		Meanings Subsection		
Meanings conveyed by words, phrases, statements, and reading selections	17 items	8 items	7 items	32 items
Meanings derived from relationships among words, phrases, and statements		2 items	3 items	5 items
		Techniques, Structures, and Forms Subsection		
Techniques used to present meaning, and the meaning or effect created by the use of allusion, analogy, dialect, euphemism, irony, malapropism, metaphor, repetition	5 items	9 items	3 items	17 items
Structures and forms used to present meaning such as genre, plot, foreshadowing, narrative point of view, transition		5 items	2 items	7 items
		Presentation of Character Subsection		
Presentation of character, narrator, or author through reference to attitudes, traits, motives, types, relationships		5 items	14 items	19 items
TOTAL	22 items	29 items	29 items	80 items

Table 3

English 33 Item Distribution

PART B: MULTIPLE-CHOICE

English 33 Achievement Test June 1982	Knowledge and Understanding	Application and Synthesis	Appreciation and Evaluation	Total
		Meanings Subsection		
Meanings conveyed by words, phrases, statements, and reading selections	15 items	6 items	1 item	22 items
Meanings derived from relationships among words, phrases, and statements	4 items	6 items	6 items	16 items
		Techniques, Structures, and Forms Subsection		
Techniques used to present meaning, and the meaning or effect created by the use of analogy, irony, parody, alliteration, metaphor, pun, allusion, paradox, spoonerism, symbol	3 items	11 items	5 items	19 items
Structures and forms used to present meaning such as poetic form, foreshadowing, narrative point of view, modes of discourse (description, narration, argument, exposition)	2 items	3 items	1 item	6 items
		Presentation of Character Subsection		
Presentation of character, narrator, or author through reference to attitudes, values, traits, motives, tone	2 items	4 items	11 items	17 items
TOTAL	**26 items**	**30 items**	**24 items**	**80 items**

Determination of the Student Population

Each school jurisdiction was given the option of testing all English 30 and 33 students or testing only randomly selected classes. To ensure a very small sampling error for the jurisdiction results, however, it was necessary for small jurisdictions to test almost every class. For example, a jurisdiction with 30 classes in a given course would need to test 28 of them to reach the desired confidence level of ±2% with 99% confidence. Boards were required to notify the Student Evaluation Branch of their wish to have all of their students write. The two largest jurisdictions, Calgary School District #19 and Edmonton School District #7, opted for sampling. All other jurisdictions required all of their students to write the tests.

Complete lists of English 30 and English 33 classes being taught in the Edmonton school district and the Calgary school district at the time of the test were obtained, and random sampling techniques were used to select classes to be tested. The classes selected represented about 63% of the English 30 classes and 77% of the eligible English 33 classes in Edmonton School District #7, and 62% and 73% of those classes in Calgary School District #19.

It should be noted that in many jurisdictions, high school instruction is carried out under the semester system. Only students who were enrolled in either English 30 or English 33 in June, 1982, were eligible to be tested. Those students completing either course in the first semester were not eligible to write.

Administration

Jurisdictions were requested in April to report the number of students enrolled in English 30 and the number of students enrolled in English 33 in each school. In May, letters were sent by the Student Evaluation Branch to superintendents, principals, and teachers in the province requesting their

co-operation in the testing. Information addressed to the superintendents and principals included the test schedule, procedures for test administration, and requirements for returning test materials. Information addressed to the teachers related to the administration of the test and the return of test materials. Each jurisdiction was sent the appropriate number of tests and administration instructions, packaged according to school. After the test was administered, teachers were instructed to collect all test booklets and answer sheets and return them to the principal for forwarding to school board offices, which, in turn, were responsible for sending the test booklets and answer sheets to the Student Evaluation Branch.

Data Collection

A total of 11 506 students in 218 schools from 104 jurisdictions returned scorable Part A booklets; 7490 students completed Part B of the English 30 test and 3865 students completed Part B of the English 33 test. Of the English 30 students, 77 completed only Part B; of the English 33 students, 74 completed only Part B. Three hundred and two students completed Part A only.

Test Design Statistics

Standard test statistics indicate that the multiple-choice sections of both tests have levels of reliability suitable for providing data to compare achievement between a jurisdiction and the province. The Kuder Richardson Formula 20 reliability coefficients were .856 for English 30 multiple-choice and .835 for English 33 multiple-choice.

Chapter III

SCORING OF WRITTEN EXPRESSION

Scoring centres were established in Grande Prairie, Edmonton, Red Deer, Calgary, and Lethbridge. Each of these centres had a project leader, an administrative assistant, and either one or two head readers. The project leader exercised overall supervision, the administrative assistant was responsible for managing the flow of work and for the remuneration of markers, and the head reader trained and supervised 25 to 30 markers. Compositions were marked by more than 200 experienced teachers of high school English under the direction of Alberta Education staff. If markers experienced difficulties or had questions, they consulted the head reader, who either dealt with the issue or notified the project leader.

Before papers were sent to the scoring centres, student identification information was removed and the compositions were sorted according to topic and placed in bundles of 15. When marking any composition, the scorers had no way of knowing the identity of the student who wrote the composition or the course in which the student was enrolled.

The first morning of the scoring session was used for training those markers who were to serve as small group leaders for the consistency checks, which were held for the purpose of maintaining reliability of scoring. The first afternoon was used to train the markers.

Consistency checks were held both in training sessions and once each morning and afternoon throughout the marking sessions. At these checks, all markers were given identical sets of four papers. They independently assigned total expression scores and a score for one aspect of writing to each of four papers. Scoring guides used by the markers are given in Figure 5. The group leaders scored total expression and all four aspects for each paper. Group

membership changed at every check. After scoring, group members compared results, and the person giving the highest score for each paper explained the strengths in the composition that had led to that score. The group leaders recorded all scores.

The scoring guide consisted of a set of five descriptors for each of five marking scales. The scales were as follows: Total Expression, Thought and Detail, Organization, Matters of Choice, and Matters of Convention. Descriptors in each set were assigned values from 1 (the lowest level of achievement) to 5 (the highest). Figure 5 presents the scoring guides and descriptors used by the markers.

All papers were given four readings. The administrative assistant would distribute a bundle of 15 papers to each marker. The marker would then read and score each of the papers for total expression and one aspect. The first time a paper was read, the particular aspect scored would be Thought and Detail; the second time, Organization; the third time, Matters of Choice; and the fourth time, Matters of Convention. The marker would record the assigned marks and his or her own three-digit identifying number on a detachable tab on the back cover that was preprinted with the booklet serial number. After scoring a bundle of papers, the marker would return it to the administrative assistant. The tab for each booklet would be removed and the marker number recorded on the booklet. Before issuing the bundle for the next reading, administrative staff would check marker numbers to ensure that no papers were scored twice by the same marker.

The total expression score was considered the most important aspect to be evaluated. If at least three of the four scores for Total Expression were identical, the composition was assigned that score. In the event of discrepancy, the paper was scored by a fifth reader. The median or central point of the five assessments was used in such cases. The score for Total Expression, then, was based on the pooled judgment of four or five markers.

Figure 5

The Written Expression Scoring Guide

SCORING OF WRITTEN EXPRESSION: TOTAL EXPRESSION

Writing will be scored on the strengths shown in the selection, organization, development, and expression of the ideas and feelings you present on the topic chosen.

CATEGORIES

5 Some writing gives the impression that it is the product of ***exceptional*** thought and expression. Carefully chosen and closely related supporting thoughts and details develop the central idea or theme. The content is organized in such a way that the writer's interpretation, attitude, and purpose are unquestionably evident throughout the work. The expression reveals a mature use of language structures, spelling, and vocabulary. Such writing demands respect.

4 Some writing gives the impression that it is the product of ***proficient*** thought and expression. Appropriate supporting thoughts and details develop the central idea or theme. The content is organized in such a way that the writer's interpretation, attitude, and purpose are clearly seen. The expression reveals a mastery of language structures, spelling, and vocabulary. Such writing prompts attention.

3 Some writing gives the impression that it is the product of ***competent*** thought and expression. Conventional thoughts and details develop the central idea or theme. The content is organized in such a way that the writer's interpretation, attitude, and purpose are generally clear. The expression reveals an awareness of and attention to language structures, spelling, and vocabulary. Such writing is usually accepted as adequate.

2 Some writing gives the impression that it is the product of ***limited*** thought and expression. Casual thoughts and details suggest but only partially develop the central idea or theme. The content is presented in such a way that the writer's interpretation, attitude, and purpose may be unclear. The expression reveals a lack of assurance in language structures, spelling, and vocabulary. Such writing communicates without fully satisfying the reader and possibly without satisfying the writer.

1 Some writing is ***unacceptable*** as the product of thought and expression at the Grade 12 level. A central purpose or idea is neither explicitly developed nor implied. Detail is non-supportive or non-existent. Content is disorganized or presented so as to leave in doubt the writer's interpretation, attitude, and purpose. The expression reveals faults in language structures, spelling, and vocabulary which impede communication and leave the reader dissatisfied.

ASPECT CATEGORIES

THOUGHT AND DETAIL

SCORE	DESCRIPTION
5	Ideas are refreshing and invite the reader to read with anticipation.
4	Ideas are appropriate to purpose and have some claim to uniqueness.
3	Ideas are acceptable but commonplace.
2	Haphazard selection of ideas sometimes obscures purpose.
1	Paucity of ideas: inadequate to convey purpose or point of view.

ORGANIZATION

SCORE	DESCRIPTION
5	Work is skillfully crafted: ordered, coherent, and with apt emphasis throughout.
4	Work is proficiently organized: organized and coherent throughout most of the work.
3	Work is organized with considerable attention to order and coherence.
2	Work presents haphazard organization with occasional evidence of order and coherence.
1	Work shows inadequate sense of order and coherence.

MATTERS OF CHOICE

SCORE	DESCRIPTION
5	Sophisticated selection of words and use of structures result in accuracy, precision, and effectiveness.
4	Proficient selection of words and use of structures contribute to accurate and appropriate expression.
3	Competent selection of words and use of structures indicate considerable attention to accuracy and precision.
2	Limited selection of words and use of structures indicate questionable attention to accuracy and precision.
1	Inadequate selection of words and use of structures indicate insufficient attention to accuracy and precision.

MATTERS OF CONVENTION

SCORE	DESCRIPTION

Awareness of spelling, punctuation, capitalization, and format demonstrates:

5	Sophisticated handling of conventions.
4	Proficient handling of conventions.
3	Considerable attention to conventions.
2	Sporadic attention to conventions.
1	Lack of knowledge or willingness to attend to conventions.

Chapter IV

RESULTS

Guide to Interpretation

The results of the English 30 and English 33 Achievement Tests are most useful when used to compare achievement in a particular jurisdiction with the provincial targets. Care, however, must be exercised in making these comparisons.

Many factors influence the performance of students within a jurisdiction. This makes setting provincial targets a difficult task. Among the many factors that influence the results in English 30 and English 33 are the following:

1. The tests have been designed to sample the English 30 and English 33 curricula. The amount of testing time limits the extent to which it is possible to test the objectives in the curriculum.

2. Whenever achievement is assessed, there are errors of measurement. These arise from many sources and can never be entirely eliminated. Care has been taken to develop scoring procedures with a high degree of reliability, but perfect accuracy is not possible.

3. The size of the jurisdiction relates to the stability of results from one year to the next. Average levels of ability are much more likely to fluctuate in smaller jurisdictions than in larger ones, with the consequence that the average achievement scores attained by smaller jurisdictions are likely to fluctuate as well.

4. Longer tests are more reliable than shorter tests. This means that longer subtests give more accurate indications than shorter ones. It implies that since the sample of writing is limited to one product, the assessment may not be as accurate as if several essays were written and scored.

5. The motivation levels of the students greatly influence the achievement scores.

In setting realistic targets for achievement, further consideration had to be given to the nature of the tasks performed and to the two populations that took the tests (English 30 and English 33). Given the factors mentioned previously, and given the care taken in developing and monitoring the scoring system, it was decided to use the central category as the basis for setting provincial targets in the Part A - Written Expression section of the test.

For the purposes of interpreting the results, it was judged reasonable to aim to have 80% of the English 30 students score in the top three categories. For English 33 students, a target of 50% was judged reasonable. For the multiple-choice portion of the English 30 test, the target mean was set at 55%, or an average raw score of 44; for English 33, the target mean was set at 45%, or an average raw score of 36. The lower target mean for English 33 takes into consideration the difficulty level of the test and the extent to which English 33 students were able to respond to the items throughout the province. These target means were selected after a thorough examination of the results and after careful consideration of the difficulty level of the tests. They are intended to serve as an indication of a reasonable level of achievement for a typical jurisdiction.

As noted earlier, jurisdictional averages are subject to various influences that tend to reduce their accuracy. To accommodate this, target means were translated into regions or performance bands so that individual jurisdictions could assess their own performances. For the multiple-choice tests and their subtests, these regions were set at five percentage points around the target mean.

The essay test targets are expressed as percentages of students receiving a score of 3 or higher. The regions were set at 10 percentage points around the target. This broader region takes into consideration the fact that the scoring was based on judgments required in marking the compositions, and the fact that only five levels of scoring were used.

Target Regions for English 30 and 33 - Written Expression

	Score Range	Target Region (% of students)
English 30	3 - 5	70 - 90
English 33	3 - 5	40 - 60

Target Regions for English 30 and 33 - Multiple-Choice

	Number of Items	Target Region (Average Raw Score)
English 30:		
Total	80	40 - 48
Meanings	37	18 - 22
Techniques, Structures, and Forms	24	12 - 14
Presentation of Character	19	9 - 11
English 33:		
Total	80	32 - 40
Meanings	38	15 - 19
Techniques, Structures, and Forms	25	10 - 12
Presentation of Character	17	7 - 9

To interpret the results for a jurisdiction, the reader should first examine the test specifications, read the sample questions and the scoring categories, and decide whether the target means are appropriate for that jurisdiction. For jurisdictions with fewer than 25 students writing, the targets should be accepted with caution.

Next the jurisdiction averages on the multiple-choice component should be compared with the target regions and areas of strength and weakness noted. Because the provincial values are of interest, the reader may wish to compare performance with the provincial values. For the total score on the multiple-choice test, jurisdictional averages within four raw score points of the provincial average would indicate no reliable difference between the jurisdiction and the provincial average. For the subtests, two points would indicate a reliable difference.

When comparing the Written Expression results with the target regions, the reader should keep in mind that the values being compared are proportions of students, not average scores. These provide only one indication of jurisdictional performance, and the reader should also look at the proportion of Category 5 performances that were received from students, as well as the proportion of Category 1 performances. (See Tables 4, 5, 6, 7, and 8.) No guidelines have been set for Categories 1 and 5, but the provincial results provide a standard for comparison. For these results a 5% difference is an adequate rule of thumb to use in assessing reliable differences from the provincial results.

Absentees

If more than 10% of the eligible students in a jurisdiction did not write the test, the reported average for that jurisdiction may not accurately represent the jurisdiction averages. Teacher-assigned marks for students who did not write could be compared with teacher-assigned marks for students who did write. If the averages are the same for the two groups, the reported jurisdiction averages are probably representative. If the averages are different, some estimates can be made of what the jurisdiction averages might have been if all students had written the test. Jurisdictions with high absentee rates may wish to contact the Student Evaluation Branch for assistance in estimating their jurisdiction averages.

In examining the test results, the reader must keep in mind that a test score does not indicate why a particular performance occurred, but only that it did occur. After studying the results, the identification of reasons for

that performance should be undertaken. There are a variety of factors that should be examined:

1. Student motivation. Consideration should be given to the degree to which students were motivated to perform to their levels of ability.

2. Student ability. While the notion of a target region is designed to take into consideration year-to-year fluctuations in the average ability levels of students, it is possible that a group of students with a particularly high or low average ability may come through a system. This is much more likely to be a factor in small systems than in large ones.

3. Teaching and curriculum. Consideration should be given to the type of instruction students have received in the jurisdiction and the adequacy of curricular implementation.

There will be other factors that are of importance in particular jurisdictions. School boards wishing to examine further the results in the light of local factors are encouraged to establish their own local interpretation panels.

Part A - Written Expression Results

Tables 4 to 8 on the following pages give the frequency distribution of scores for Part A. These tables should be interpreted using the category descriptions given in the preceding chapter.

Table 4

Frequencies of Total Expression Scores

Score Category	English 30 Frequency	%	English 33 Frequency	%	Combined Frequency	%
5	92	1.2	1	0.0	93	0.8
4	1259	16.8	41	1.1	1300	11.5
3	4845	64.7	1342	35.3	6187	54.8
2	1224	16.4	2124	55.9	3348	29.7
1	65	0.9	294	7.7	359	3.2

Table 5

Frequencies of Thought and Detail Scores

Score Category	English 30 Frequency	%	English 33 Frequency	%	Combined Frequency	%
5	517	6.9	28	0.7	545	4.8
4	1941	25.9	294	7.7	2235	19.8
3	3272	43.7	1385	36.4	4657	41.3
2	1510	20.2	1520	40.0	3030	26.8
1	245	3.3	575	15.1	820	7.3

Table 6

Frequencies of Organization Scores

Score Category	English 30 Frequency	%	English 33 Frequency	%	Combined Frequency	%
5	426	5.7	23	0.6	449	4.0
4	1737	23.2	204	5.4	1941	17.2
3	3322	44.4	1310	34.5	4632	41.0
2	1787	23.9	1745	45.9	3532	31.3
1	213	2.8	520	13.7	733	6.5

Table 7

Frequencies of Matters of Choice Scores

Score Category	English 30 Frequency	%	English 33 Frequency	%	Combined Frequency	%
5	339	4.5	8	0.2	347	3.1
4	1432	19.1	118	3.1	1560	13.8
3	3821	51.0	1313	34.5	5134	45.4
2	1743	23.3	1926	50.7	3669	32.5
1	150	2.0	437	11.5	587	5.2

Table 8

Frequencies of Matters of Convention Scores

Score Category	English 30 Frequency	%	English 33 Frequency	%	Combined Frequency	%
5	261	3.5	3	0.1	264	2.3
4	1799	24.0	182	4.8	1981	17.6
3	3606	48.2	1423	37.4	5029	44.6
2	1521	21.7	1694	44.6	3315	29.4
1	198	2.6	500	13.2	698	6.2

Tables 4 to 8 indicate that the typical level of achievement in written expression is lower among English 33 students than among English 30 students. Of the separate aspects evaluated in both courses, the highest typical level of achievement was in Thought and Detail, and the lowest was in Matters of Choice.

Part B - Multiple-Choice Results

Table 9 gives the provincial results for the full tests and for subtests.

Table 10 shows the provincial results for the different cells and rows of the English 30 Part B blueprint. These norms were calculated from a weighted frequency distribution: the results for the two sampled jurisdictions were weighted by the proportion of students sampled in each jurisdiction.

Table 9

English 30 Multiple-Choice Results
(Raw Scores)

	Mean	Standard Deviation
Full Test (80 items)	43.5	10.5
Subtest: Meanings (37 items)	21.0	5.4
Techniques, Structures, and Forms (24 items)	12.2	3.7
Character (19 items)	10.4	3.0

Table 10

Provincial Norms (Means and Standard Deviations)
English 30

MULTIPLE-CHOICE SECTION

English 30 Achievement Test June 1982	Knowledge and Understanding	Application and Synthesis	Appreciation and Evaluation	Total
		Meanings Subsection		
Meanings conveyed by words, phrases, statements, and reading selections	10.6* (2.7)	4.0 (1.7)	3.6 (1.4)	18.3 (4.8)
Meanings derived from relationships among words, phrases, and statements		0.9 (0.7)	1.8 (0.9)	2.7 (1.2)
		Techniques, Structures, and Forms Subsection		
Techniques used to present meaning, and the meaning or effect created by the use of such techniques as allusion, analogy, dialect, euphemism, irony, malapropism, metaphor, repetition	2.7 (1.3)	4.4 (1.8)	1.3 (0.9)	8.4 (2.9)
Structures and forms used to present meaning, such as genre, plot, foreshadowing, narrative point of view, transition		2.8 (1.2)	0.9 (0.7)	3.8 (1.5)
		Presentation of Character Subsection		
Presentation of character, narrator, or author through reference to attitudes, traits, motives, types, relationships		2.9 (1.2)	7.4 (2.3)	10.4 (3.0)

*Each cell contains the raw score mean and raw score standard deviation for that cell. In the first cell, for example, the raw score mean is 10.6 and the raw score standard deviation is (2.7).

Tables 11 and 12 give the results for English 33, using the same procedures as for English 30.

Table 11

English 33 Multiple-Choice Results (Raw Scores)

	Mean	Standard Deviation
Full Test (80 items)	34.9	9.8
Subtest: Meanings (38 items)	17.2	5.3
Techniques, Structures, and Forms (25 items)	10.2	3.3
Character (17 items)	7.4	2.8

Table 12

Provincial Norms (Means and Standard Deviations)
English 33

MULTIPLE-CHOICE SECTION

English 33 Achievement Test June 1982	Knowledge and Understanding	Application and Synthesis	Appreciation and Evaluation	Total
		Meanings Subsection		
Meanings conveyed by words, phrases, statements, and reading selections	7.5* (2.6)	2.5 (1.4)	0.3 (0.4)	10.3 (3.4)
Meanings derived from relationships among words, phrases, and statements	1.9 (1.0)	2.7 (1.3)	2.3 (1.3)	6.9 (2.6)
		Techniques, Structures, and Forms Subsection		
Techniques used to present meaning, and the meaning or effect created by the use of such techniques as analogy, irony, parody alliteration, metaphor, pun, allusion, paradox, spoonerism, symbol	1.5 (0.8)	3.9 (1.8)	2.1 (1.1)	7.5 (2.6)
Structures and forms used to present meaning, such as poetic form, foreshadowing, narrative point of view, modes of discourse (description, narration, argument, exposition)	0.8 (0.7)	1.4 (0.9)	0.6 (0.5)	2.7 (1.3)
		Presentation of Character Subsection		
Presentation of character, narrator, or author through reference to attitudes, values, traits, motives, tone	1.1 (0.7)	1.5 (1.0)	4.8 (2.1)	7.4 (2.8)

*Each cell contains the raw score mean and raw score standard deviation for that cell. In the first cell, for example, the raw score mean is 7.5 and the raw score standard deviation is (2.6).

Tables 9 and 10 show performance on the multiple-choice part of the English 30 test. Although performance in different parts of the test shows some variation, these scores are not directly comparable. The sets of items that make up the different parts were not selected to be equal in average level of difficulty. Thus, differences may be due to variations in item difficulty rather than in student performance. In combination with jurisdictional results, however, the norms can be used to detect patterns of relative strength or weakness in achievement.

Tables 11 and 12 show performance on the multiple-choice part of the English 33 test. The restrictions for interpretation mentioned for English 30 also apply to English 33.

It must be emphasized that the purpose of the Achievement Testing Program is to produce results valid at the jurisdictional and provincial levels, not at the individual class or student level. Because of restrictions of time and space, the test questions are only a small sample of possible questions for any curricular objective. Thus, individual student scores may vary greatly, depending on the specific questions selected. Accordingly, the achievement tests were not designed to evaluate individual students, but rather to evaluate the performance of programs within a jurisdiction or within the province. In addition, conclusions about the performance of any given class should not be made because of the small number of students in most classes and because there is a large variation in ability of students in different classes.

For large groups of students, variations due to question selection will average out. If the numbers are sufficient, it is meaningful to compare the achievement of groups of students to the provincial averages. Jurisdiction results are presented in tables which parallel the tables of provincial results to facilitate comparisons.

Chapter V

CONCLUDING OBSERVATIONS

In this report, guidelines are provided to assist local boards and administrators in interpreting the achievement test results for their jurisdictions. Target regions are designed to help local decision-makers assess whether the achievement of their students, when compared to the provincial results, is sufficiently different to be educationally important.

The provincial results relate to the target means and target regions in the following ways:

The provincial mean of 54% on the multiple-choice section of the English 30 test is very close to the target mean of 55% and well within the target region of 50% to 60%.

The results indicated that the multiple-choice section of the English 33 test was more difficult than anticipated for English 33 students. The target mean was adjusted to reflect this for purposes of interpreting the results. Thus, the provincial mean of 44% falls well within the adjusted target region of 40% to 50%.

The frequency of written expression scores in the three highest categories for English 30 is within the target region of 70% to 90%.

The frequency of written expression scores in the three highest categories for English 33 is 37%, which is below the target mean of 50% and below the target region of 40% to 60%.

The following limitations of the tests and, therefore, of the results should be borne in mind:

The findings are limited to those that can be obtained from a pencil-and-paper test; many skills developed in language arts cannot be measured by this type of test.

Although speaking, listening, and viewing are important aspects of the language arts curriculum, the test was limited to reading and writing.

The items on the test cover a representative sample of the reading and writing objectives in the language arts curriculum.

The test is based on the curriculum prescribed by the *Program of Studies*. It is not possible to reflect each local jurisdiction's unique interpretation of the program.

Nonetheless, both English 30 and English 33 tests were found to be reliable and valid instruments. Standard statistical practices indicate that the multiple-choice sections of both tests have suitable levels of reliability. Content validity was ensured by careful adherence to the curriculum specifications and by involvement of external validation committees.